

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



QuadR-Tree Indexing Selection Engine for Tuning Spatial Database System using Mobile Geographical Information System Technology

¹Maruto Masserie Sardadi, ²Daut Daman, ³Mohd Shafry and ⁴Zahabidin Jupri
*Faculty of Computer Science and Information System
 University Technology of Malaysia,
 81310 Skudai, Malaysia*

1. Introduction

The latest Mobile GIS technology is probable to manage the spatial components of daily business project, in the corporate database, and to apply proper geographic analysis efficiently in a desktop-focused application. This technology uses wireless internet for transferring process of spatial data from server to Mobile GIS or vice versa. However, the problem in wireless Internet is system bottlenecks that can slow down in the process of transferring data. One of the reasons why system bottlenecks could happen in this situation is because the spatial data is too large. The database design therefore is the most important part that must be considered. Database systems are now expected to be self-managing, self healing, and always-up. Another parts which also contribute in this problem are network, application server and web server (Kwan & Shi, 2002).

Most of previous researchers have experimented to change the structure of R-tree and Quadtree spatial data indexing method that could give better performance. However, it still leaves some issues and problems that need to solving. One of the most common issues and problems is the method only applicable to some of applications such as data arrangement, but it needs speed improvement in the transfer process and data retrieval. This chapter gives different methodology that concentrates on fine tuning of a spatial database system using R-tree and Quadtree spatial data indexing method, without changing the structure of those spatial data indexing method.

This chapter introduces the creation of indexing engine selection to tune the spatial database system. This indexing engine selection, as it call QuadRtree Engine Selection, uses Rule Base – Knowledge Base Expert System (KBES) to select between R-tree and Quadtree spatial data indexing method. Those spatial data indexing methods have different advantages and disadvantages based on the condition of the data. This engineering innovation is one of the challenges in actual global technology. The objectives of this chapter are to make an optimization in transferring data and thus to provide the latest accurate information for mobile users with a limited communication bandwidth.

2. R-tree Spatial Data Indexing

R-trees are tree data structures that derived from B-trees. It is used for representing multidimensional point data, but is used for spatial process methods for indexing multi-dimensional information. Each node in the tree communicates to the smallest d-dimensional rectangle that surrounds its son nodes. The leaf nodes include cursor to the real geometric entity in the database, as an alternative of childs. The entity are represented by the smallest associated rectangle in which they are enclosed (Guttman, 1984).

Frequently the nodes communicate to disk leaf, and therefore the constraint defining the tree is selected so that a small amount of nodes is appointed during a spatial query. Note that rectangles matches to different nodes may be related. Also a rectangle may be spatially controlled in numerous nodes; so far it can be correlated with only one node. This indicates that a spatial query may regularly necessitate a number of nodes to be visited before discovering the existence or absence of a particular rectangle.

R-tree spatial data indexing method uses Minimum Bounding Rectangle (MBR) in the algorithm method. The following figure denotes the rectangle part of an index entry E by EI, and the tuple-identifier or child-pointer part by EP. Give an R-tree whose root node is T, and find all index records whose rectangles overlap a search for rectangle S.

- S1 [Search sub trees] If T is not a leaf, check each entry E to determine whether EI Overlaps S. For all overlapping entries, invoke Search from the tree whose root node is pointed to by Ep.
 - S2 [Search leaf node] If T is a leaf, check all entries E to determine whether EI overlaps S. If so, E is a qualifying record.

Fig. 1. R-tree algorithm (Guttman, 1974)

In Mobile GIS technology, R-tree has been chose for further explanation in this chapter. The reason for choosing R-tree is that the application dealing with data containing and several overlap like what can be seen in Mobile GIS technology will bring better results, and the application dealing with data containing "whole Earth objects", bounding-box based methods will not work properly. In addition, Mobile GIS technology, will store, retrieve and process spatial data. Spatial data in GIS is a mandatory thing that needs to be solved for most all of the problems that might come.

3. Quadtree Spatial Data Indexing

A Quadtree is a tree data composition used to exploit a set hierarchical data compositions. The common property is that they are based on the principle of recursive decomposition in which each internal node has up to four children. Generally, Quadtree separate two dimension space through divide it into four sub part or region. The part could be in rectangular, arbitrary or square shape. Raphael Finkel and J.L. Bentley gave a name Quadtree for this data structure in 1974. This spatial data indexing has similar partitioning method with Q-tree. Quadtree has general decomposition approach, it decompose the space into flexible cell which has maximum capacity. The region split, then directory tree chase Quadtree spatial decomposition when meet the optimum capacity (Finkel & Bentley, 1974).

According to Spatial Concept in Oracle 10g white paper, “Quad tree spatial data indexing method uses fixed-size tiles to cover geometry. Fixed-size tiles are controlled by tile resolution. If the resolution is the sole controlling factor, then tessellation terminates when the coordinate space has been decomposed a specific number of times. Therefore, each tile is of a fixed size and shape”. Fig 2. shows the process of Quadtree.

- The quadtree begins with a square in the plane; this is the root of the quadtree.
 - This large square can be broken into four smaller squares of half the perimeter and a quarter the area each; these are the four children of the root.
 - Each child can in turn be broken into 4 sub squares to get its children, and so on.

Fig. 2. Quadtree algorithm (Finkel & Bentley, 1974)

Currently Quadtree is used for point data, curves, surfaces, areas and volumes. It decomposes to the same parts on each level, or it may be manage by the input. This peculiarity, in computer graphics is often expressed in terms of image space ladder against object space ladder. The decomposition resolution may be fixed in advance, or it may be manage by materials of input data. In some of the applications the data formation on the origin of whether they state the restrictions of sections also can be distinguished.

4. Rule-based Knowledge Base Expert System

Expert system contain a memory evoke utility and a conjecture mechanism. This formation is like a human brain. It develops the data in the same way as expert people illustrates on the skill, knowledge and judgment expanded through understanding to achieve conclusion about a problem (Chow et al., 1995). For accuracy and sensibility, the effort’s result of knowledge base expert system problem solving must be evaluated. The output should always examine as advisory (Uzel, 1987). An expert system approach process well when the problem to be solved is ill-defined or multifarious, and when decision and familiarity are helpful tools in finding the solution (Norman & Lesser, 1994).

Ruled-based systems are a comparatively simple form that may be adjusted to some of difficult problems. With Artificial Intelligence (AI), rule-based system has its strength as well as limitation that must be reflect on before deciding if it’s the appropriate procedure to use for a given problem. Generally, rule-based systems are only viable for problems in which any and all knowledge in the predicament area can be written in the outline of if-then rules and for which this problem area is not large (Zhang et al., 2008). Another advantage of using Rule-Based is that it can represent in a natural and understandable way the knowledge that summarize and explain the data.

Base on the previous exploration knowledge, it is need to setting up the weight of the rules to get the order of importance. All of the knowledge and weight has been set up in the ruled-based data set. This data set is used to decide whether it is need to use R-tree or Quadtree spatial data indexing in the Mobile GIS system. This research tries to give the weight from all of the knowledge above regarding the effect that can be affected to the system. The highest weight of the rule is the highest effect that can be affected to the system, thus the priority to chose is highest. Some of the rules are compulsory to be applied without considering another rules, because these rules are depend to the condition of the data and requirement of the system, otherwise our system will perform slower. This knowledge and

dataset are dynamic, which can be updated again if the new knowledge has been explored. Here is the rule-based data set with the order of importance base on our research.

| No | R-Tree Indexing | Quadtree Indexing | Order of Importance |
|----|--|--|--------------------------------|
| 1 | An R-tree index is required for a whole-Earth index. | A quadtree index cannot be used for a whole-Earth index. | First Choice. Use R-tree |
| 2 | The approximation of geometries cannot be fine-tuned. (Spatial uses the MBR) | The approximation of geometries can be fine-tuned by setting the tiling level and number of tiles. | Second Choice. Use Quadtree |
| 3 | You can index up to four dimensions. | You can index only two dimensions. If LRS (Linear Referencing System) data is indexed using a spatial quadtree index, only the first two dimensions are indexed; the measure dimension and its values are not indexed. | Third Choice. Use R-tree |
| 4 | Heavy update activity to the spatial column may decrease the R-tree index performance until the index is rebuilt. | Heavy update activity does not affect the performance of a quadtree index. | Weight:15. Use Quadtree |
| 5 | If your application workload includes nearest-neighbor queries (SDO_NN operator), R-tree indexes are faster, and you can use the sdo_batch_size keyword. | If your application workload includes nearest-neighbor queries (SDO_NN operator), quadtree indexes are slower, and you cannot use the sdo_batch_size keyword. | Weight: 13. Use R-tree |
| 6 | An R-tree index is recommended for indexing geodetic data if SDO_WITHIN_DISTANCE queries will be used on it. | A quadtree index is not recommended for Indexing geodetic data if SDO_WITHIN_DISTANCE queries will be used on it. | Weight: 9. Use R-tree |
| 7 | Less storage is required. | More storage is required. | Weight: 4. Use R-tree |
| 8 | Index creation and tuning are easier. | Tuning is more complex, and setting the appropriate tuning parameter values can affect performance significantly. | Weight: 3. Use R-tree |

Table 1. R-tree and Quadtree Rule-based data set

The knowledge expert from R-tree and Quadtree Rule-based data set above has been applied to geographical map, which is Pasir Gudang Map (one of the location in Malaysia). This map was used as GIS data case. It is needs to apply the Rule-Base data set to all of the tables in Pasir Gudang database. It also need exploration to recognized whether those tables suitable to use the R-tree, Quadtree or without spatial data indexing method.

5. QuadRtree Engine Selection

QuadR-tree Engine Selection is a technique of selecting an R-tree and Quadtree spatial data indexing method in terms of data condition. There are some considerations when choosing R-tree or Quadtree spatial data indexing. The approach is how to make a decision to choose between R-tree or Quad-tree spatial data indexing that is suitable for different cases of geographical data based on spatial data indexing knowledge. Thus there would be one or two spatial data indexing method in one single spatial database system. The following figure is the algorithm of QuadRtree engine selection.

QuadRtree spatial data indexing Engine Selection is based on rule-based, Knowledge Base Expert System (KBES). The rule-based approach contains all of the proper encoded knowledge; generally written in If-Then rules. It also puts in working memory which holds some data or information. Firstly it needs to explore the knowledge base of R-tree and Quadtree spatial data indexing with the advantages and disadvantages comparison of those spatial data indexing methods. Here is the QuadRtree Engine Selection algorithm:

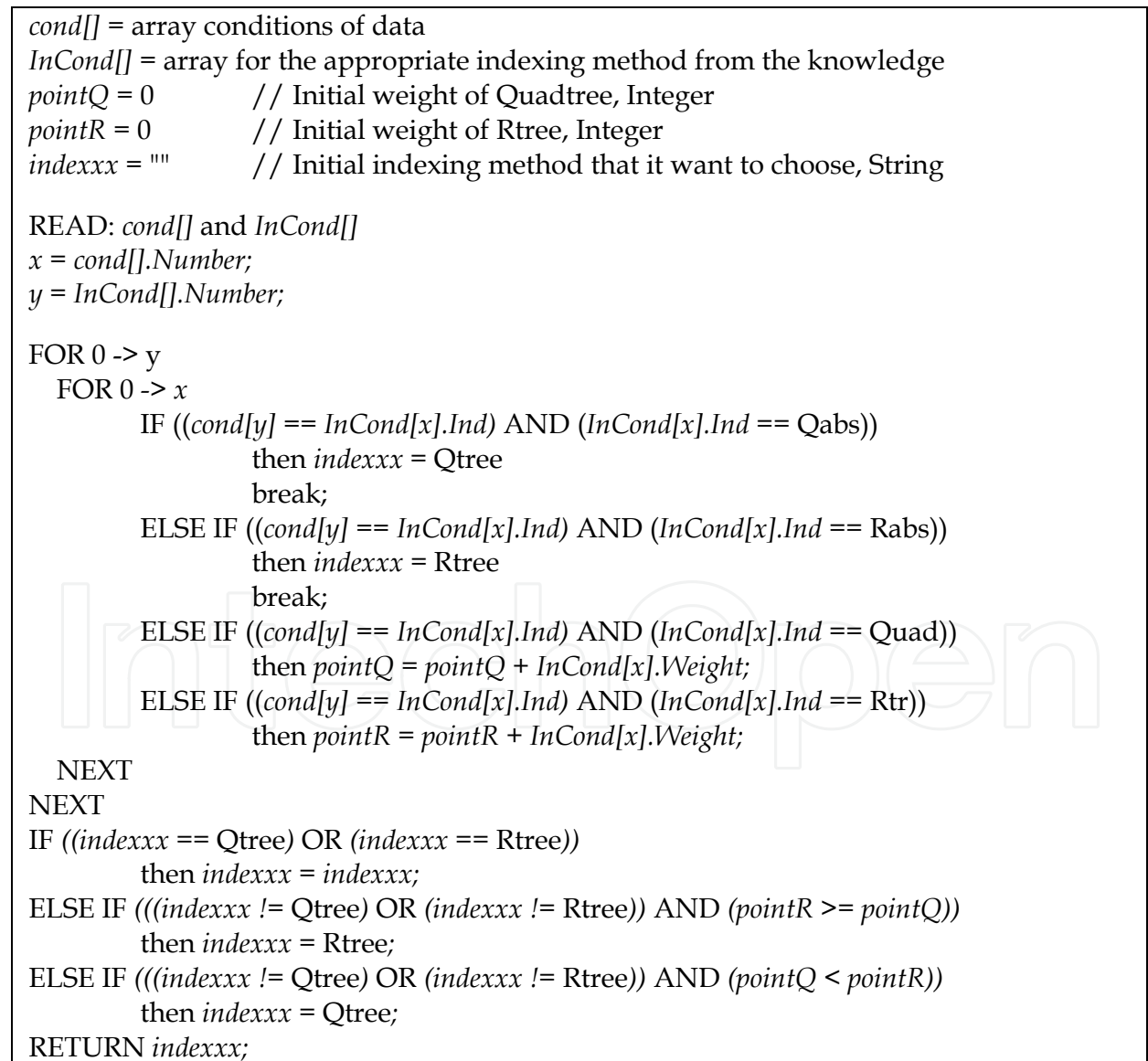


Fig. 3. QuadRtree selection engine algorithm

All of the exploration of knowledge and weight would be put in the InCond[] array as it called rule set. This rule set will be used to compare with the input data cond[] array to match and get the appropriate spatial data indexing method. Through a set of assertions, which collectively form the 'working memory', and a set of rules that specify how to act on the assertion set, a rule-based system will be created. This rule set is dynamic; it could be added or updated again from the new knowledge's exploration of those spatial data indexing comparisons.

Before reading the set of data input, it is important to create the weight initialization for R-tree (pointR) and Quadtree (pointQ) spatial data indexing method that will be used in the selection engine's algorithm. These pointQ and pointR are used to decide whether R-tree or Quadtree spatial data indexing method which will be used based on the weight of data condition (cond[]). The initial indexing method (indexxx) also needs to create that be change to the appropriate spatial data indexing method.

The loop will be created to read and match all of the input data conditions cond[] and the rule set provided knowledge InCond[]. There are two methods in the InCond[] array they are InCond[].Ind to get the condition from the rule set knowledge and InCond[x].Weight to get the priority weight. Next, that there are two conditions from the knowledge they are absolute condition and the priority condition. The absolute condition is the condition that should use the proposed indexing method; R-tree (Rabs) or Quadtree (Qabs) without concerning to another condition and knowledge. If there is more than one similar proposed indexing method, it can use R-tree or Quadtree spatial data indexing method. The priority condition is the condition that needs to calculate the proposed amount of the weight from R-tree (Rtr) and Quadtree (Quad) spatial data indexing method. Thus it should choose the highest result of the weight amount either from R-tree or Quadtree spatial data indexing method.

6. Testing & Evaluation

In order to evaluate the selection engine algorithm, this chapter presents the result of Mobile GIS application development. The selection engine algorithm has been putted in that application and could measures the speed of geographical map process over Mobile GIS network. The proposed indexing selection engine algorithm has tested using Pasir Gudang map. The developed Mobile GIS application has common GIS function, but it has additional function which is searching feature to measure the selection engine algorithm performance. The model used for testing and evaluation in this research has been simplified with the following assumptions:

- One dimension time
- Oracle 10g spatial databases are used to develop the GIS Map data;
- Hyper Text Transfer Protocol (HTTP) utilized with data transfer protocol was used for connection from server to client or vice versa. Data retrieval process uses an active TCP/IP connection with providing mobile user initial position.

Simplified mobile GIS equipment has demonstrated how process retrieval and transfer data works. A mobile user access data's from server that uses Oracle 10g Spatial Extension as a SDBMS to show the geographical map. The following way is to characterize the performance of this function; The average response time from the mobile client is measured as the time spent (in seconds) from the moment of the query is issued to the moment of the

query is generated. The equipments and applications used in this research are summarized in the following table.

| | | |
|-------------------------------------|-------------|---|
| Web Server & GIS Server: | CPU | Intel Pentium Core 2 Duo |
| | RAM | 2 GHz |
| | Web Server | IIS |
| | GIS Server | Oracle Application Server, MapViewer and MapGuide Open Source |
| | Database | Oracle 10g Spatial |
| Mobile Device Equipment: | Model | Tablet PC |
| | OS | Windows XP |
| | Web Browser | Internet Explorer |

Table 2. Setting of the experiments

This chapter tried to evaluate the geographical data using R-tree, Quadtree, QuadRtree and without using spatial data indexing method. This evaluation has been tried to Pasir Gudang Map. The total number of records used in the spatial database is around 35.000. Some steps are needed to measure the retrieve and transfer processing time. It was tested from the number of record 1000, 5000, 10000, 15000, 20000, 25000, 30000 until 35000. Here is the evaluation process to Pasir Gudang geographical maps.

The coordinate system used in the geographical map is XY-M (Arbitrary X-Y Coordinates (Meters)) with lower left coordinates X: 37,478.171281 Y: -7,091.537339 and upper right coordinates X: 39,341.018451 Y: -66,250.168600. The initial view of the map is center X: 38,409.594866 Y: -66,670.852970 with zoom to 1: 32,965.970000. Our map has also built some of the general GIS map functions, like: Select Mode, Clear Selection, Zoom in, Zoom Out, Help, Pan Mode, Buffer, Select Polygon, Select Radius, Select Within, Measure, Print, and Search.

The Pasir Gudang Map spatial database use layers to organize the map and make easier to access. These layers also used to develop the map order which pins down one layer to another. There are four layers used in this spatial database. The first one is Default Layer which has BANGUNAN_ELLIPSE, BANGUNAN_POLYLINE, BANGUNAN_REGION, BANGUNAN_TEXT, MINIMUM_TROUGHPUT_REGION, and SEMPADAN_KAWASAN_REGION tables. The second one is LOT Layer which has LOT_LUAR_REGION and LOT_REGION tables. The third one is JALAN Layer which has JALAN_POLYLINE table. And the last layer is ZONING Layer which has ZONING_REGION table. The order for all of those layers is start from Default Layer, LOT Layer, JALAN Layer and ZONING Layer. The order of layers will take the effect in the layers position.

Figure 4. show the process of searching data or location. This searching method can choose the data or location by LOT, ZONING or REGION. The QuadR-tree indexing engine selection run in this process. Start from find the data or location until the data will be shown and zoom like in Figure 5. Furthermore, this process measure the speed of using indexing or not over Mobile GIS network.

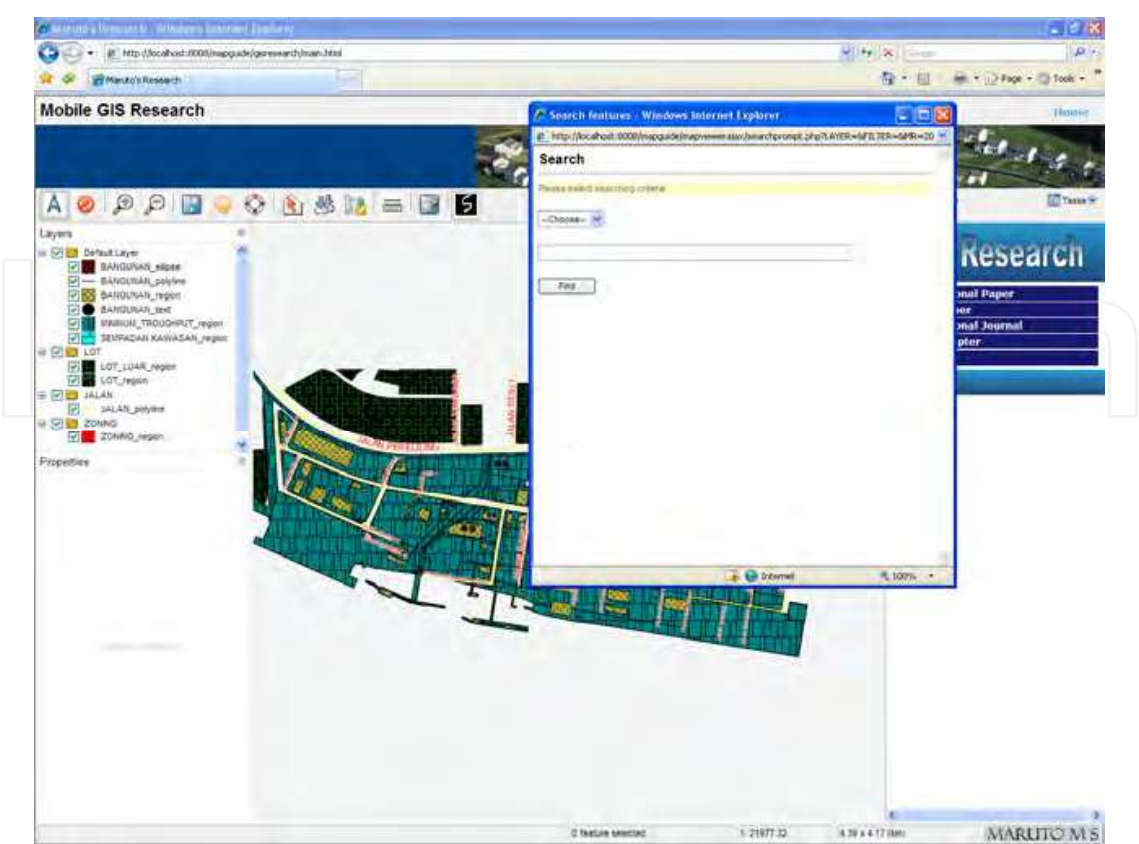


Fig. 4. Search the data or location.

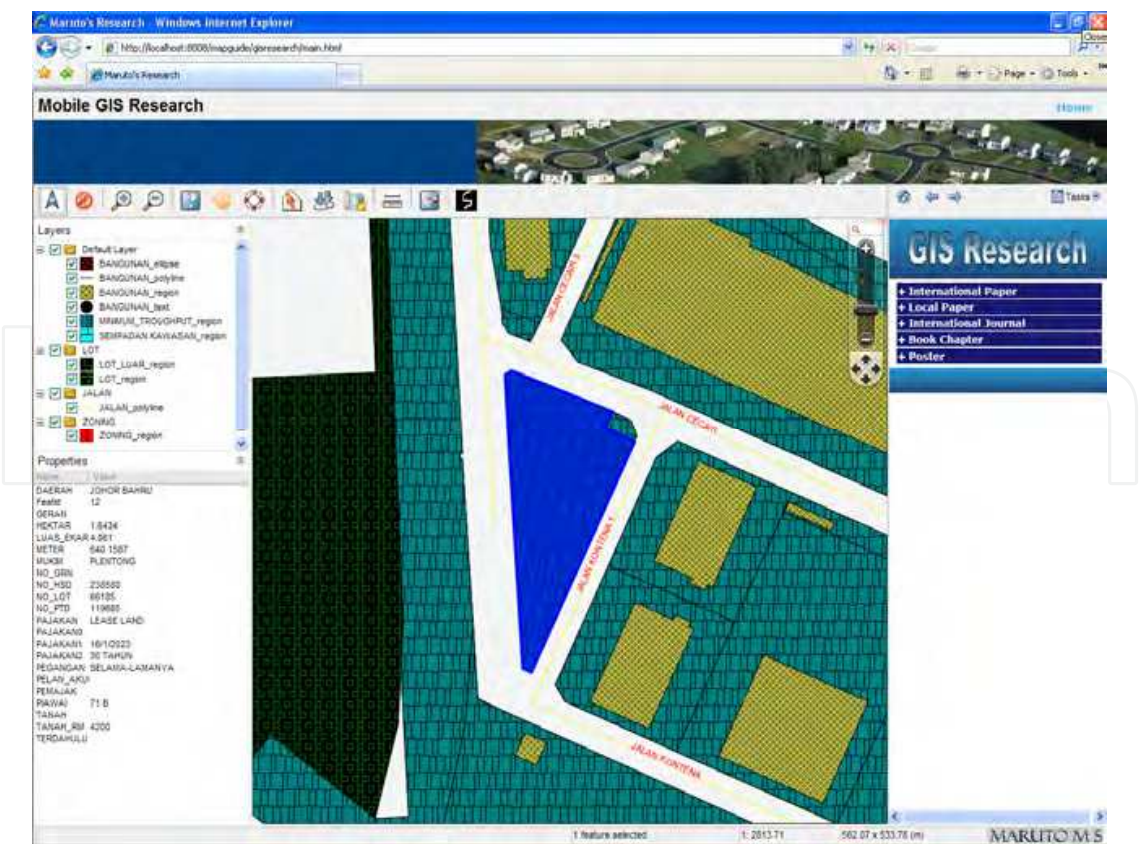


Fig. 5. Zoom the searched data.

The Pasir Gudang Map application result is depicted in Figure 6., which is plotted as a two-dimensional graph to illustrate the results of average response time using Mobile GIS application. The line with square point corresponds to the time measurement from the data use an R-tree spatial data indexing. The time measurement for the data that use Quadtree spatial data indexing has plotted with circle point. While the line with triangle point represents the time measurement for the data that doesn't use spatial data indexing. This research observes that the time saved increases as the database size expands. For the number of records, over 30,000, the time required in the data that uses an R-tree as well as Quadtree spatial data indexing are almost two times more than the data that doesn't use spatial data indexing. It also can be seen from the result, the used of QuadRtree indexing selection engine algorithm can make faster the process. The QuadRtree indexing is plotted in the line with plus point. It is noted, however, that this is related to the actual time of the query, since the database volume is varied as the record changes. The response time is slightly different to the result shown that figure. Nevertheless, it can easily discover that the response time on the data that uses an R-tree as well as Quadtree spatial indexing are faster than the data that doesn't use spatial indexing, moreover if the system use QuadRtree indexing selection engine algorithm.

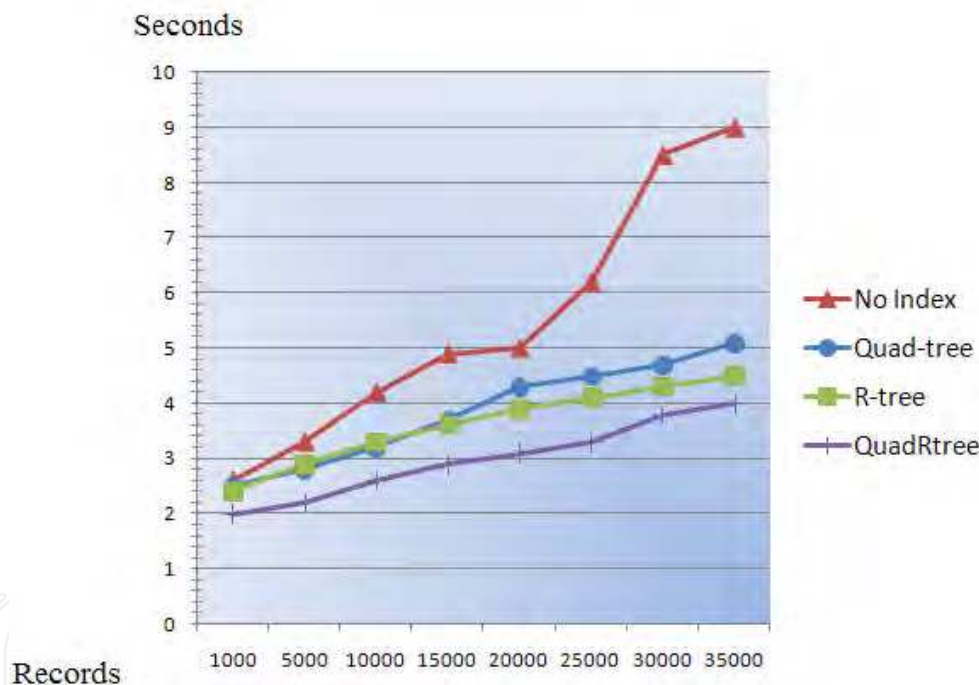


Fig. 6. The average response time.

7. Conclusion

The use of QuadR-tree indexing engine selection can improve the speed of transfer and retrieve data over Mobile GIS network. However, this engine still not automatically input the spatial data and condition of the data. This research also can be extended to explore the new knowledge of R-tree and Quadtree spatial data indexing that will be putted in the dynamic Rule-based data set. This knowledge will be used for selection consideration in QuadR-tree indexing engine selection algorithm. The Knowledge could be advantages or disadvantages of those spatial data indexing method.

8. References

- Chow, P. K. O., Yeung, D. S. (1995). Multidimensional Knowledge Structure. *Expert Systems with Applications*. Vol. 9, No. 2, 1995, pp. 177-187.
- Finkel, R., & Bentley, J. L. (1974). Quad Trees: A Data Structure for Retrieval on Composite Keys. *Acta Informatica* 4 (1): 1- 9.
- Guttman, A. (1984). R-Trees: A Dynamic Index Structure for Spatial Searching. Proc. 1984 ACM SIGMOD International Conference on Management of Data, pp. 47- 57.
- Kwan, K., & Shi, W. (2002). *A Study of Dynamic Database in Mobile GIS*.
- Norman, C. & Lesser, V. (1994). The Evolution of Blackboard Control Architectures. *Expert Systems with Applications*, Vol. 7, No. 1, 1994, pp. 1-30.
- Oracle Spatial 10g White Paper (2006). *Oracle Spatial Quadtree Indexing, 10g Release 1*.
- Uzel, A. R. (1987). *Guidelines for Expert System Applications*. CEM, Feb. 1987, pp. 40-45.
- Zhang, L. & Xi, L. F. (2007). A Novel Fractal Image Coding Based on Quadtree Partition of the Adaptive Threshold Value. *Theoretical Advances and Applications of Fuzzy Logic and Soft Computing*, 504-512.

IntechOpen



New Advanced Technologies

Edited by Aleksandar Lazinica

ISBN 978-953-307-067-4

Hard cover, 350 pages

Publisher InTech

Published online 01, March, 2010

Published in print edition March, 2010

This book collects original and innovative research studies concerning advanced technologies in a very wide range of applications. The book is compiled of 22 chapters written by researchers from different areas and different parts of the world. The book will therefore have an international readership of a wide spectrum.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Maruto Masserie Sardadi, Daut Daman, Mohd Shafry and Zahabidin Jupri (2010). QuadR-Tree Indexing Selection Engine for Tuning Spatial Database System using Mobile Geographical Information System Technology, New Advanced Technologies, Aleksandar Lazinica (Ed.), ISBN: 978-953-307-067-4, InTech, Available from: <http://www.intechopen.com/books/new-advanced-technologies/quadr-tree-indexing-selection-engine-for-tuning-spatial-database-system-using-mobile-geographical-in>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2010 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen